tendency to be associated with dissolved or particulate/sediment fractions, assuming that other, non-contaminant related factors (*e.g.*, TOC, AVS) are similar in concentration:

- Dissolved fraction components - salts of nickel, copper, cadmium, and dichromate have a tendency to be more prevalent in the dissolved phase than those of other metals when conditions permit;

- Particulate/sedimentary fraction components - PAHs, PCBs, Dioxins, DDE, silver, lead, zinc, arsenic, mercury, and trivalent chromium have a greater tendency to be particle-associated than the above group, again assuming comparable geochemical conditions.

This information will be used to aid in the interpretation of contaminant distribution, bioaccumulation and toxicity as discussed in the following sections.

5.2. Toxicity Evaluations

Site-specific evaluations of bulk surface sediments were conducted using the 10-day amphipod (*Ampelisca abdita*) mortality test. The bulk sediment test is a tool used to evaluate the bioavailability of contaminants in the respective media. Comparison of these results to trophic transfer effects measurements (discussed in Section 5.3) will provide a detailed spatial evaluation of potential impacts to aquatic biota.

The amphipod *Ampelisca abdita* was used to determine potential impacts of sediment contaminants from the Raymark study areas.

Background. The euryhaline benthic amphipod, *Ampelisca abdita*, which ranges from Newfoundland to Florida and the Gulf of Mexico, was used to evaluate the toxicity of sediments from Raymark. This tube-dwelling amphipod constructs a soft, upright, membranous tube 3 to 4 cm long in fine-grained sediments from the intertidal zone to a depth of 60 m. *Ampelisca* ingest either surface-deposited particles or particles in suspension, and respire in both overlying and interstitial waters.

The 10-day amphipod test has been used extensively to assess the toxicity to *Ampelisca abdita* of laboratory-spiked and field collected sediments (DiToro *et al.*, 1992; Scott and Redmond, 1989; and Long *et al.*, 1990). In addition, *Ampelisca abdita* has been used routinely for sediment toxicity tests conducted by SAIC in support of numerous EPA programs (SAIC, 1990a; SAIC, 1991; SAIC, 1992a; and SAIC, 1993a). *Ampelisca abdita* was the most sensitive species tested in the U.S. EPA/U.S. ACE Field Verification Program, and has represented the toxicological basis for EPA research on the availability of metals in relation to acid volatile sulfides in marine sediments (Gentile *et al.*, 1987 and DiToro *et al.*, 1992). *Ampelisca abdita* has also been used to characterize the toxicity of sediments from the Calcasieu River, LA,

which cover a broad range of salinity and grain size (SAIC, 1990b). It was the first species used to demonstrate the toxicity of sediments from New Bedford Harbor, MA, and subsequently was used to assess the effectiveness of capping procedures as part of a Pilot Dredging Project on site remediation techniques (U.S. ACE, 1989). SAIC has recently conducted toxicity tests of sediments from New York Harbor for EPA Region II and the New York District (SAIC, 1992b; SAIC, 1994a; and SAIC, 1995), as well as a series of tests for NOAA which characterize toxicity of sediments from the Hudson-Raritan Estuary, Long Island Sound, Boston Harbor, and Tampa Bay (SAIC, 1992c; SAIC, 1992d; SAIC, 1993b; and SAIC, 1994b).

Methodology. Amphipod tests (5 replicates each) were conducted on surface sediments from 16 intertidal stations at Raymark in Stratford, CT. Amphipods were exposed to test sediments for 10 days under static conditions, following SOPs developed according to ASTM and EPA procedures (ASTM, 1990 and U.S. EPA, 1994; Appendix B-1). Water quality parameters monitored daily throughout the test included temperature, salinity, dissolved oxygen (DO), and pH. Five replicates were performed for each sample. In addition, samples were analyzed for total ammonia in order to assess any potential toxic effects of ammonia under the static test conditions. Porewater ammonia was measured on days -1 and 0; and ammonia in the overlying water column was measured on days 5 and 10. Sediments were press-sieved and homogenized before placement into the test chambers, after which, sub-samples were collected for porewater ammonia analyses.

Performance control sediments were collected from the amphipod collection site sediments in the Narrow River, RI. Survival of *A. abdita* using collection site sediments from the Narrow River has ranged between 90 and 100% (n = 18) with a mean of 98.1% over a 5 year period.

Data analyses. Stations with a mean survival less than that of the performance control were compared statistically to the control using a one-way, un-paired *t*-test (alpha=0.05) assuming unequal variance. An examination of a large historical data set has shown that *Ampelisca abdita* percentage survival data meet the requirement of normality, therefore data were not transformed. Threshold for evidence of toxicity for *Ampelisca abdita* has been defined as survival statistically less than the performance control and ≤ 80% of the mean control survival. Statistical power curves created from SAIC's extensive testing database with *A. abdita* show that the power to detect a 20% difference from the control is approximately 90%. Sites meeting both requirements (statistically different than the performance control and survival ≤ 80% of the control) were flagged.

Results. Results of the sediment assay are summarized in Table 5.2-1. Mean performance control survival (not shown) was 92%. Mean sample survival, normalized to mean performance control survival, ranged between 3.3 to 82.6%. Mean amphipod survival for 15 stations was statistically different from survival in the performance control sediment and was < 80% of mean performance control survival. Complete data are presented in Appendix B-1.

Total ammonia concentrations in sediment and water (porewater) are often a primary source of interference to the interpretation of CoC impacts as sample toxicity. Total ammonia values during the amphipod test ranged from 0.30 to 4.67 mg/L. Un-ionized ammonia concentrations were not calculated in this study because concentrations of total ammonia were low, thus un-ionized ammonia concentrations were expected to be low, as well. Figure 5.2-1 illustrates, amphipod survival versus total ammonia for the Raymark sample stations. Reduced amphipod survival at Stations C-3 and D-6 does not appear to be caused by total ammonia concentrations. In fact, some Raymark stations with increased ammonia concentrations actually have increased survival (Figure 5.2-1).

5.3. Existing Toxicity-Based Criteria and Standards

Toxicity-based criteria and standards provide the basis for comparing expected or actual environmental concentrations of contaminants to toxicological benchmark concentrations, thereby allowing an estimation or quantification of potential adverse effects. For the present risk assessment, the primary benchmarks utilized were 1) ER-L/ER-M values presented in Long et al. (1995); 2) EPA water quality screening values derived from Water Quality Criteria (WQC), as presently available or as predicted from sediment-based benchmarks and partitioning parameters. Discussion of each of these benchmarks and its relation to the present risk assessment is included below.

ER-L/ER-M values. The NOAA ER-L and ER-M concentrations correspond to the lower 10th and 50th percentiles, respectively, of all concentrations of a contaminant observed to cause a biological effect, over a range of studies and species (Long and Morgan, 1990; Long et al., 1995). Conceptually, ER-Ls are similar to LOELs (lowest observed effect levels), which represent the lowest toxicant concentration observed in bioassays to cause biological effects. Another type of benchmark, called the Apparent Effects Threshold or AET (PTI 1988; U.S. EPA, 1989d), was also evaluated to address individual contaminants in sediments and represents the level of individual chemicals above which statistically significant biological effects are always expected to occur. As discussed in Section 3, the ER-L values are typically more conservative (*i.e.*, corresponding to lower benchmark levels) than AET values, usually representing concentrations that are an order of magnitude lower. As part of the conservative approach taken for the present assessment, ER-L/ER-Ms were used as the primary sediment-based benchmarks.

Water Quality Criteria. The U.S. EPA Water Quality Criteria (WQC) were used to evaluate the toxicological significance of CoCs measured in porewater samples collected in the present investigation. Water-quality criteria are based on the total recoverable concentration of the metal as sampled from test chambers during aquatic toxicity tests. However, it is recognized that the dissolved metal concentration more closely approximates the bioavailable fraction of the metal in the water column than does the total extractable concentration. Lussier and Boothman (1995) addressed this

issue through paired (unfed and fed) toxicity tests with *Mysidopsis bahia, Mysidopsis lateralis* and *Ampelisca abdita*. Data from these tests were used to derive conversion factors between the dissolved phase effect concentration of metals and the total recoverable concentration of metals. For most metals (arsenic, cadmium, dichromate, lead, nickel, selenium, and zinc), the conversion factor was 0.95 or greater, indicating that the presence of food (which could potentially alter metal bioavailability) had minimal effect on the derived Water Quality Criteria. One exception was copper, for which the conversion factor (0.83) indicated that the expected dissolved phase concentration is 83% of the promulgated criteria. These findings suggest that WQC can be used to evaluate CoC concentrations in pore water, although organically enriched conditions in the field sample may modify chemical bioavailability.

The U.S. EPA 1998 WQC were used to evaluate contaminant concentrations in porewater and elutriates. Formerly, the U.S. EPA 1986 WQC were used. A comparison of the two criteria shows the 1998 criteria does not provide benchmark values for organics. Rather, sediment based benchmark values were used to calculate water quality screening values for organics.

Changes were apparent between the 1986 and 1998 benchmark values for some metals. Upon examining the U.S. EPA 1986 and 1998 benchmark criteria saltwater acute (WQC-SA) values, only copper, mercury and silver were different by >10%. The 1998 criterion for copper was 65% higher than the criterion developed in 1986. The WQC-SA benchmark value for mercury changed by a factor of 14.29% between 1986 and 1997, while the WQC-SA benchmark for silver decreased from 7.2 μg/L (U.S. EPA, 1986) to 1.9 μg/L (U.S. EPA, 1998b), a 74% change. For the saltwater chronic criteria (WQC-SC) values, none of the analytes had values that differed between the U.S. EPA 1986 values and the 1998 values by more than 10%, with the exception of mercury. It is of interest to note that U.S. EPA's new criteria value for mercury (0.94 μg/L) is a substantial departure from the 1986 value (0.025 μg/L).

Sediment Quality Criteria. For CoCs lacking WQC values, a WQC-equivalent concentration can be derived, in some cases, based on the conversion of the sediment-based benchmark into water based units assuming equilibrium partitioning (discussed further in Section 6). For non-ionic organic substances, Sediment Quality Criteria (SQC) are derived directly from the product of the Water Quality Criteria (WQC) and the organic carbon partition coefficient (K_{oc}) of the chemical (Di Toro *et al.*, 1991). Here it is seen that the SQC approach uses the WQC as the appropriate effects concentration for protection of organisms in sediment, even though the WQC were derived from water column species. The primary uncertainties with this approach are: 1) whether the pathway of chemical exposure (*e.g.*, water column, interstitial water or ingested sediment) under equilibrium conditions is an important determinant in chemical bioavailability, and 2) if organism habitat or feeding behavior may result in more limited exposure to non-ionic organic contaminants (*e.g.*, epibenthic or filter-feeding organisms receive significantly less exposure than infaunal species).

Recent research has found that differences in organic CoC bioaccumulation are

not apparent among benthic species despite differing trophic mode
(Tracey and Hansen, 1996). The present assessment also provided suitable data for
demonstration of this relationship (Section 6.1). This observation reduces the
uncertainties identified above with regard to the applicability of the EqP approach to
derive criteria for ecological effects assessment in the present investigation.

5.4. Uncertainty

Numerous assumptions are made concerning the applicability of 1) toxicity
evaluations and 2) particular benchmarks used as criteria and/or standards to evaluate
impacts to biota. These assumptions bear upon the certainty of risk derived from these
effects-based measures; i.e., whether the observed response falsely suggests an effect
where none exists (a "false positive") or alternatively, fails to detect an effect where one
is occurring (a "false negative"). An assessment of uncertainty in each area follows.

Toxicity evaluations. The evaluation of ecological effects of contaminated
sediments using toxicity tests is essential because chemical concentrations alone are
not accurate predictors of biological effects. The principal advantage of the sediment
toxicity testing approach is that the tests are performed in a manner comparable to
WQC derivation exercises (e.g., mortality or sublethal effects are observed), hence the
data are directly comparable to these criteria. Uncertainties associated with toxicity
testing conducted in the present study are that the responses may not be chemical-
specific, and the responses observed may not represent chronic effects. In addition,
there is uncertainty in the comparability between the sediment test species and the
water test species upon which the WQC are derived (U.S. EPA, 1998b).

Sediment Benchmarks. As summarized in Section 5.4, the derivation of
ER-L/ER-Ms is based on very conservative assumptions concerning use of the lower
10th and 50th percentiles, respectively, of all concentrations of a contaminant that have
been observed to cause biological effects. In the derivation of screening criteria
(Section 3.3), ER-L values in particular are generally lower by one order of magnitude
for most parameters (including AET values, representing the only other effects-based
benchmark that is commonly applied), and hence were most often the benchmark of
lowest value for each CoC. The uncertainty is the level of conservatism that is
appropriate to assess potential ecological risk. ER-Ls are used in this assessment to
provide a protective evaluation; however, these benchmarks may be overly
conservative as they do not account for site-specific factors that can mitigate (buffer)
the responses of ecological systems to particular contaminants. The incorporation of
toxicity data at various effects levels and for species from different phyla and trophic
levels is an attempt to add another measure of realism to the final assessment. A
further issue with the use of the ER-L/ER-Ms is the somewhat limited list of chemicals
for which values are available compared to the overall list of contaminants from the
study, leaving the possible effects of some contaminants not fully assessed.

Use of Surrogate/Indicator Species. The species evaluated in this study,

including ribbed mussels (*Modiolus demissus*) and one bioassay organism (*Ampelisca*) were used as indicators of the assumed general response of the various communities within the study region. These species represent a variety of biological endpoints which have been shown to be sensitive to contaminant inputs and whose relationship to a particular habitat and community is well established. It also was deemed important to maintain consistency in the use of the same species, for purposes of data comparability, among the various phases of the risk assessment. Nonetheless, the use of surrogate or indicator species is associated with some undefined level of uncertainty, since one or a few species cannot precisely duplicate the response of the numerous species that comprise the various communities of the region.
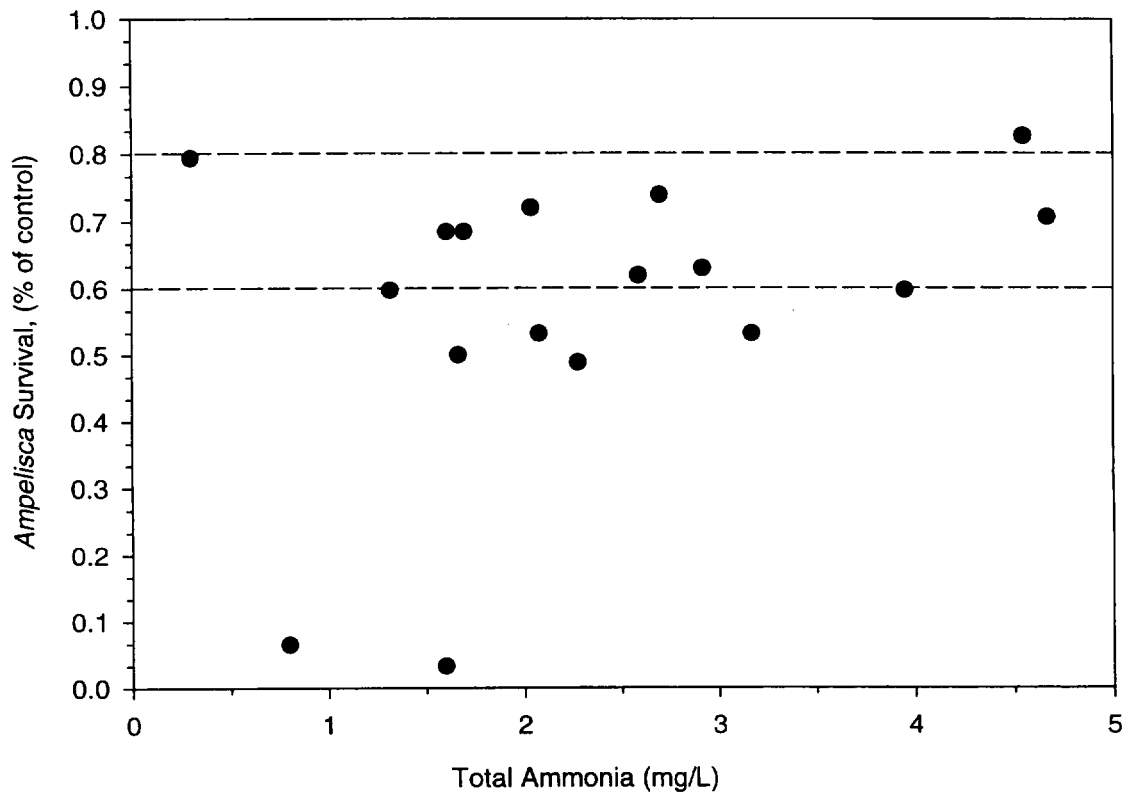
Figure 5.2-1. Amphipod (*Ampelisca*) survival vs. Total Ammonia measured
in sediment porewaters from the Raymark study area. The horizontal
dashed lines indicate interpretive threshold values for slight (< 80% of control)
and severe (< 60% of control) impact on amphipod survival.

Table 5.2-1. Summary of toxicity test results using *Ampelisca* survival for the Raymark Phase III Ecological Risk Assessment Investigation.

Sediment Toxicity

| Sample ID | 10-Day Solid-Phase Amphipod Survival Test | | |
| | *Ampelisca abdita* | | |
	Total Ammonia (mg/L)	Survival Mean[1]	Effects Ranking[3]
C-1	3.17	53.3 *	++
C-2	2.08	53.3 *	++
C-3	0.80	6.5 *	+++
D-1	0.30	79.3 *	+
D-2	1.32	59.8 *	++
D-3	1.67	50.0 *	++
D-4	2.59	62.0 *	+
D-5	1.61	68.5 *	+
D-6	1.60	3.3 *	+++
E-1	1.70	68.5 *	+
E-2	2.92	63.0 *	+
E-3	2.70	73.9 *	+
E-4	2.28	48.9 *	++
F-1	3.95	59.8 *	++
F-2	4.55	82.6 -	-
F-3	4.67	70.7 *	+
Reference[2]	2.04	72.0	+

1 - Percent *Ampelisca* survival in bulk sediment sample normalized to laboratory performance control sediment. *Ampelisca* survival: - = no effect; * = sample survival was statistically different and < 80% of performance control survival.
2 - Reference = GM08 (SAIC, 1998).
3 - Effects Ranking:
  + = statistically < control and 60-80% of control;
  ++ = statistically < control and < 60% control; and
  +++ = statistically < control and < 20% control.

6.0. RISK CHARACTERIZATION

In the Risk Characterization phase of the Raymark ERA, the following activities have been conducted: (1) Comparisons of exposure point concentrations against established standards and criteria for sediments and water quality (Section 6.1); (2) Assessment of tissue residue exposure and effects in target receptors (Section 6.2); (3) An analysis of CoC bioaccumulation/trophic transfer in birds and mammals (Section 6.3); (4) Analysis of toxicity versus CoC concentrations (Section 6.4); (5) Risk Synthesis (Section 6.5); and (6) Analysis of Uncertainties associated with the above interpretations (Section 6.6).

An important element of risk characterization is the communication of study results in an easily discernable fashion. In studies such as the present ERA where multiple Weights of Evidence (WoE) are used to characterize chemical exposure and biological effects at dozens of stations, the volume of numeric data can be formidable. Hence, a systematic, semi-quantitative summary of results is generated for each WoE by applying exposure and effects rankings so as to facilitate the assessment of potential risk (discussed in Section 6.6). Accordingly, Tables 6.0-1 and 6.0-2 are presented to summarize the ranking strategies used for estimated chemical exposure (Exposure Ranking) and biological effects (Effects ranking). The WoE-specific criteria are discussed in detail in their respective sections. The application of the ranking criteria results in four tiers of probability of adverse exposure or effects; baseline ("-") , low ("+"), intermediate ("++") and high ("+++"). It is acknowledged that the selected ranking criteria are somewhat subjective and are largely based on best professional judgement. However, they represent a synthesis of the data, and a concerted effort was made to provide a comparable and consistent approach across various WoE.

6.1. Comparison of CoC Concentrations with Criteria and Standards

In this section, concentrations of contaminants of concern (CoC) are compared with effects-based screening benchmarks. For sediments, comparisons were made against the NOAA ER-L and ER-M values (Long and Morgan, 1990 and Long et al., 1995), whereas porewater concentrations were compared against EPA Water Quality Criteria. For the contaminants in each matrix for which there were benchmarks available, Hazard Quotients were developed by dividing the contaminant concentration measured at the station by the benchmark concentration.

6.1.1. Bulk Sediment Contaminants

Calculated ER-L and ER-M Hazard Quotients for sediments are presented in Appendix Tables D-1-1 and D-1-2, respectively. Table 6-1.1 presents a visual summary of exposure-based weights of evidence for Raymark surface sediment contaminants according to the evaluation criteria described in Table 6.0-1. Total PCB HQs for Area C stations fell between the ER-L and ER-M benchmark values. Total PCBs at these three

stations exceeded the ER-L benchmark (22.7 ng/g; Long *et al.*, 1995, Appendix D-1-1). Stations in Area D had Total PCB HQs ranging from baseline at Station D-1 to high at Station D-5. Low exposure to Total PCBs was apparent at Stations D-2, D-3, and D-6, and an intermediate exposure was observed at Station D-4 (Table 6.1-1). The highest concentrations of PCBs in sediments were observed in Areas E and F. Three stations in Area E had concentrations exceeding twice the ER-M (180 ng/g, Long *et al.*, 1990, Appendix Table D-1-2). In one instance, an extremely high ER-M HQ of 232 was calculated for Total PCBs (Station E-1). Two Area F stations (F-2 and F-3) had Total PCB HQs greater than twice the ER-M value, while Station F-1 had an intermediate exposure for Total PCBs. Total PCB concentrations at the reference station only exceeded the ER-L (Table 6.1-1), producing a low hazard quotient ranking for that station.

Exposure-based weights of evidence for PAHs measured in sediments in the Raymark study area are summarized in Table 6.1-1. In Area C, low exposures were observed for most PAHs at Station C-2, including HMW, LMW and Total PAHs. The remaining two stations had baseline exposure for all PAHs. In Area D, baseline exposures were observed for Stations D-1 and D-2. However, some of the highest exposures due to PAHs were observed at Station D-3, with most concentrations two- to four-fold higher than the ER-M (Table 6.1-1; Appendix Table D-1-2). Stations D-4 and D-5 had low exposures for most PAH compounds, while Stations D-6 had baseline exposure for eleven PAHs (including LMW, HMW and Total PAHs) and low exposure for three. Stations in Area E had baseline or low exposure for PAHs with the exception of Station E-3. HQs at this station ranged from above the ER-L to slightly more than two times the ER-M for benzo(a)anthracene and pyrene. HQs for Stations in Area F were highest at Station F-3 with PAH concentrations being two- to seven-fold higher than the ER-M for all PAHs except acenaphthylene and fluorene. High exposures were also apparent for three PAHs (benzo(a)pyrene, dibenz(a,h)anthracene, and pyrene) at Station F-2. Seven PAHs at Station F-2 had intermediate exposures, including LMW and HMW PAHs; the remaining four PAH exposures were low. The reference station had one intermediate exposure for dibenz(a,h)anthracene, but the remainder of the exposures were either low or baseline.

HQs were generally below the ER-M benchmark for pesticides. In Area C, only p,p'-DDD had a low exposure at Station C-2. Low exposures for p,p'-DDD were also apparent in Area D at Stations D-4, D-5 and D-6 (Table 6.1-1). In Area E, low exposures were apparent for all four pesticides at Stations E-2 and E-4. Station E-3 had a low exposure for p,p'-DDE. The remainder of the pesticides had low exposures at Stations E-1 and E-3. HQs for pesticides in Area F ranged from below the ER-L to above twice the ER-M. At Station F-1, an intermediate exposure to p,p'-DDD was apparent. The remainder of the exposures for this site were either low or baseline. Station F-2 had a high exposure for p,p'-DDD, with an HQ marginally greater than two times the ER-M. High exposures were also evident for p,p'-DDD and p,p'-DDE at Station F-3, (HQ = 4.4 and 3.7, respectively; Appendix D-1-2). Reference data were not available for comparison against sediment benchmarks.

The majority of stations had baseline exposure concentrations for dioxins (2,3,7,8 - TCDD equivalent concentrations). Exceptions were noted at Station D-5 and E-1, with an HQ of 2.23 (Appendix D-1-2). The HQ for dioxins at Station E-1 was also greater than twice the ER-M (Table 6.1-1). In Area F, dioxin exposures were baseline for all three sites, and baseline values were also apparent at the reference station.

In summary, the above results illustrate the patchy distribution of organic contaminants in surface sediments from the Raymark study area. Within the study areas, the highest ER-M HQs were observed for Total PAHs in Area F (0.5 to 2.9); Station D-3 along the Housatonic River also exhibited very high exposure (HQ = 1.6). Exposures to Total PCBs were highest at Stations E-1 and D-5, with HQs of 232 and 28, respectively. Dioxin exposure was highest at Station D-5 near the Housatonic River, with a Dioxin-Fish ER-M HQ of 2.23.

Concentrations of metals in sediments from the Raymark study area relative to the ER-L and ER-M benchmarks are summarized in Table 6.1-1; Hazard Quotients are presented in Appendix D-1-1 and D-1-2. In Area C, lead and zinc exceeded ER-L benchmark values at Station C-2 and C-3. ER-L benchmarks were also exceeded for copper at Station C-1 and mercury at Station C-1 and C-2. Intermediate exposure was evident given that copper exceeded ER-M benchmark values at Stations C-2 and C-3, while mercury concentrations were above the ER-M at Station C-3. None of the stations had high exposure for metals.

Concentrations of metals in Area D had either low or baseline exposures. Station D-4 had exposures above the ER-L benchmark value for all metals except cadmium.

In Area E, copper, lead and mercury exceeded the ER-M at Station E-1. ER-L benchmarks for these metals were exceeded at the remaining stations in Area E. Chromium, nickel and zinc concentrations also exceeded ER-L benchmarks at Station E-1. None of the Area E stations had high exposure to metals.

The highest exposure for metals was apparent at Area F, Station F-2. HQs for copper, lead and zinc at this location were all greater than twice the ER-M benchmark value (2.82, 2.62 and 2.40, respectively; Appendix D-1-2). In contrast, concentrations of copper, lead, mercury and zinc only exceeded the ER-L benchmark at Station F-1. Station F-3 also had low exposures for copper, mercury, nickel and zinc. Lead concentrations at this station were above the ER-M. Station F-2 had a low exposure for mercury and intermediate exposures for chromium and nickel. All metals exceeded benchmark reference values at the reference station. Copper was particularly elevated, with as HQ greater than twice the ER-M (HQ = 2.45).

Spatial patterns of anthropogenic trace metal concentrations within the Raymark study area were not pronounced. The highest exposures were apparent for copper, lead and zinc at Station F-2 on Selby Pond and copper at the reference location.

Intermediate exposures were also apparent in Area F, as well as Areas C and E. Low and baseline exposures were apparent throughout all portions of the study area.

Overall exposure rankings for sediment HQs (using criteria listed in Table 6.0-2) were low to intermediate in Area C (Table 6.1-1). In Area D, overall exposures ranged from baseline (Station D-1) to high (Stations D-3 and D-5). Intermediate and high exposures occurred in Areas E, F and the reference (Table 6.1-1). These results are brought forward to the exposure assessment summary in Table 6.1-5. No specific patterns were observed in Sediment Hazard Quotient data (Table 6.1-5). Nearly every station throughout all four areas of Raymark exceeded the ER-L or ER-M benchmark for one or more analytes. For most of these stations, the ER-L was exceeded for a number of analytes. In three cases, a single CoC exceeded the ER-M (Stations C-2, D-4, and E-4) and according to the ranking criteria, the stations were assigned to the low exposure category. This decision was supported by the fact that the CoC-specific ER-M benchmark was developed from field data where multiple CoCs (possibly at higher concentrations) could have contributed to the observed adverse effect, upon which the NOAA benchmark is based. (This assumption, however, is further evaluated by exposure-response analysis in the effects-based WoE evaluation, such that a possible CoC effect will not go undetected as a result of the above assumption). One station (D-1) had baseline exposure, since no analytes exceeded the ER-L or ER-M benchmark.

6.1.2. Sediment Porewater Contaminants

Benchmark Derivation. For this investigation, Water Quality Screening Values (WQSV) adopted primarily from EPA Water Quality Criteria - Saltwater Chronic (WQC-SC) values (or estimated equivalents, discussed below) were used as the benchmarks (Table 6.1-3). Water-based CoC criteria were derived following the decision tree presented in Figure 6.1-1. This approach allows for calculation of "WQC-SC equivalent" benchmarks, and assigns a data qualifier (DQ) to identify the benchmark source for derivation of the HQ. In Table 6.1-2, the DQ "A" is applied to benchmarks derived directly from existing WQC-SC values. For CoCs possessing WQC-saltwater acute values (WQC-SA), an 8:1 acute:chronic ratio is applied to derive the equivalent WQC-SC value (DQ = "B"). The conversion factor was derived from the mean overall acute:chronic ratio for paired chemical data contained in the EPA AQUIRE database (Shepard, 1998). Freshwater chronic data (WQC-FC) are used directly as screening values, with assigned data qualifier "C". As with WQC-SA values, freshwater acute (WQC-FA) values were converted to chronic values using an 8:1 acute:chronic ratio, and assigned DQ = "D".

Some sediment-based correlative benchmarks are required to complete the assessment of site-related CoCs where water quality benchmarks are lacking (designated as "E" in Table 6.1-2). In these cases, sediment based benchmarks (*e.g.* NOAA ER-Ls) were selected and translated into porewater equivalent concentrations using an Equilibrium Partitioning (EqP) model, described below.

Porewater based benchmarks were derived using the equilibrium partitioning (EqP) model of DiToro *et al.*, (1991). The direct applicability of Sediment Quality Criteria (SQC) has been limited by the number of available criteria to date (presently five non-ionic organic compounds including three PAHs (acenaphthene (U.S. EPA, 1993a), fluoranthene (U.S. EPA, 1993b), phenanthrene (U.S. EPA, 1993c). However, the SQC derivation process has demonstrated the applicability of WQC to porewater concentrations for prediction of sediment toxicity when partitioning characteristics of the CoC between water and the organic carbon fraction of the sediment (K_{oc}) are taken into account using the EqP model as follows:

$$1) \qquad C_p = C_s/(f_{oc} * K_{oc});$$

In the above equation, the organic chemical porewater concentration (C_p, μg/L) is calculated from the corresponding sediment concentration (C_s; μg/kg), based on the fraction of organic carbon (foc) in the site sediment (foc = %TOC/100); and the organic carbon/water partitioning coefficient (Koc) for the CoC. Values for K_{oc} (Table 6.1-3) were determined from the relationship developed by the EPA (Karickhoff, *et al.*, 1989):

$$2) \qquad \log_{10}K_{oc} = 0.00028 + 0.983*\log_{10}K_{ow};$$

where K_{ow} = the octanol/water partition coefficient.

In this process, it is assumed that the resultant value provides a level of protection equivalent to the other water quality based benchmarks used in this study. This assumption is not unreasonable given that the WQC values are designed to be protective of 95% of all species, while NOAA ER-L values represent concentrations below which 90% of all sediment samples had no measurable adverse effect. In order to derive porewater benchmark values from NOAA ER-Ls, the ER-L values were transformed into water-equivalent benchmarks using the EqP model by assuming 1% sediment TOC concentration (DQ = E). Finally, compounds for which no benchmark screening values were available are designated "NA" in Table 6.1-3.

The Water Quality Screening Values (WQSV) presented in Table 6.1-2 represent thresholds for adverse effects to aquatic biota as derived from available WQC and modified sediment benchmarks. Porewater concentrations are divided by the WQSV and WQC-SA benchmarks to obtain Porewater Hazard Quotients (PW-HQs) representing a range of exposure. These quotients are used to assess effects as discussed in this section.

Sediment Porewater Results. Table 6.1-4 presents a visual summary of exposure based weights of evidence for sediment porewater in the Raymark study area, according to the evaluation criteria described in Table 6.0-1. Calculated WQSV and WQC-SA Hazard Quotients are presented in Appendix Tables D-2-2 and D-2-3, respectively.

Total PCB HQs in Area C were all below the WQC-SC (= WQSV) benchmark value, representing baseline exposure. The same was true for Area D with the exception of Station D-5, where low exposure to Total PCBs was apparent (WQC-SC HQ 29.7, Appendix D-2-2). Station E-1 also had a low exposure to Total PCBs, with a WQC-SC HQ of 27.4 (Appendix D-2-2), while the remainder of the stations in this area had baseline exposures. All Area F stations also had baseline exposure for Total PCBs, as did the reference station.

With the exception of Station D-3, all locations in Areas C-F had low to baseline exposure. Estimated PAH concentrations in sediment porewaters were below the most conservative screening criteria for all stations in Area C (Table 6.1-4). In Area D, Stations D-1 and D-2 also had baseline exposure for all PAH concentrations. In contrast, intermediate exposure was apparent for five PAHs at Station D-3 (including Total PAHs), while high exposure was observed for HMW PAHs (WQC-SA HQ = 3.12). At the remaining stations (D-4, D-5, D-6) exposures to PAHs in porewater were low to baseline.

With a singular exception, pesticides in all areas were below the most conservative water screening criteria, indicating baseline exposure. Only a low exposure to p,p'-DDD was found at Station D-5. Reference data were not available for pesticide concentrations in sediment porewaters.

Dioxin exposures were baseline for all stations in all areas with one exception. Based on estimated concentrations of dioxins in porewater, Station D-5 had a low exposure for these constituents.

In summary, of all organic contaminants sampled, only Station D-3 along the Housatonic River had intermediate to high exposure for PAH contaminants in sediment porewaters. Overall, the porewater data suggest that organics are not a source of potential risk to aquatic biota in Areas C-F.

Concentrations of metals in surface sediment porewaters from the Raymark study area relative to the WQC-SC (= WQSV) and WQC-SA benchmarks were also assessed (Table 6.1-4). Mercury concentrations in porewater were not measured. Station D-3 had a high copper exposure, with a WQC-SA HQ of 2.33 (Appendix D-2-3). Area E also had high exposures from copper at Stations E-1 and E-2, with WQC-SA HQs of 4.33 and 3.21, respectively. An intermediate exposure to copper was apparent at Station E-3. In Area F, only Station F-2 had elevated exposure for this metal. The reference station also had high exposure for copper (HQ =11.46) as well as zinc (HQ = 4.67).

In summary, porewater exposures to metals in the study area can be characterized as low except for copper concentrations at three stations. Station F-2 on Selby Pond had an intermediate exposure for copper. However, the reference station also had high copper exposures in porewater. Stations D-3 and the reference were

placed into the high exposure category ("+++"). Two stations (E-1 and E-2) were categorized as intermediate exposure ("++"). Five stations had a low exposure ("+"). The remainder of the stations (8 stations) qualified as baseline exposure ("-") since one or no exceedence of a WQC-SA value was observed for every location (Table 6.1-4). These results are brought forward to the exposure assessment summary in Table 6.1-5.

6.1.3. Bedded Sediment Summary.

Overall exposure rankings for bedded sediment (presented in Table 6.1-5) use evaluation criteria listed in Table 6.0-2. Overall rankings are based on the above sediment and porewater HQ discussion and the SEM/AVS results presented in Section 4.3. Exposures range from baseline to high. Overall bedded sediment rankings were intermediate and low in Area C. Low exposure from bedded sediment occurred at four stations in Area D. An overall high ranking was apparent at Station D-3, due to high exposures from sediments and porewater. Overall, intermediate and low exposures occurred in Areas E, F, and the reference (Table 6.1-5).

6.2. Assessment of Tissue Residue Exposure and Effects in Target Receptors

In the section below, an analysis of CoC bioaccumulation (Section 6.2.1) in ribbed mussels as well as indicators of CoC-related exposure (Section 6.2.2) and effects (Section 6.2.3) are presented. Bioaccumulation is assessed by comparison of tissue residue concentrations with that of sediment collected at the same location. CoC exposure was assessed by comparison of site tissue residue concentrations with tissue residue concentrations from the reference station (Tissue Concentration Ratios), while effects of CoCs in ribbed mussels were addressed by comparison of tissue residues against tissue benchmarks derived from water quality criteria (Tissue Screening Concentration HQs), and comparison of tissue residues against threshold concentrations for narcotic effects (Critical Body Residue HQs).

6.2.1. Analysis of Bioaccumulation

In the sections below, the relationships between contaminant exposure and tissue residue concentration for organics (Section 6.2.1.1) and metals (Section 6.2.1.2) are discussed. Bioaccumulation of organics and metals was assessed by calculating biota-sediment accumulation factors (BSAFs) for organics and bioaccumulation factors (BAFs) for metals. These values will be used to calculate tissue concentrations of CoCs in target receptors for areas where biota collections were lacking. The following formulas were used to calculate these concentrations:

1) $[\text{Organics}]_{tissue} = [\text{BSAF} * [\text{Organics}]_{sediment} * [\text{Lipid}]] / \text{TOC}$

2) $[\text{Metals}]_{tissue} = \text{BAF} * [\text{Metals}]_{sediment}$

The following sections present the results of the calculated tissue residue-exposure relationships for target receptors in the Raymark study areas.

6.2.1.1. Analysis of Organic Contaminant Bioaccumulation

The purpose of this analysis is to identify potential organic contaminant exposure for target species representing different habitat or feeding types. For each organic contaminant class (PCBs, PAHs, pesticides, and dioxins), exposure pathway differences were evaluated through Tissue Residue - Exposure Relationships as well as Biota-Sediment Accumulation Factors (BSAFs).

Tissue Residue - Exposure Relationships. Figure 6.2-1 presents a comparison of Total PAHs, Total PCBs, and Total DDTs in ribbed mussels versus sediment concentrations of these CoCs for seventeen Raymark stations (including the reference). No significant relationships were apparent for organics data between tissue and sediment concentrations. Data for measured tissue concentrations were also examined by normalizing tissue data to lipid concentrations and sediment data to total organic carbon concentrations (Figure 6.2-2). An improved relationship was apparent, particularly for Total PCBs indicating that tissue lipid and sediment TOC are important factors governing organic contaminant bioaccumulation.

Following from the above results, Biota-Sediment Accumulation Factors (BSAFs), defined as the lipid normalized concentration of the CoC in an organism (μg/g lipid) divided by the organic carbon-normalized concentration of the same chemicals in sediment (μg/g organic carbon), were calculated to assess the bioaccumulation of organic contaminants by biota where samples were unavailable.

Results are presented for co-located ribbed mussel and sediment data in Appendix D-3-4. For the present study, the overall median BSAF values for PAHs, PCBs, and pesticides were 0.20, 2.35, and 3.88, respectively. (Tissue data for dioxins were not available, thus literature values for dioxin BSAFs were used (USEPA, 1993a)). Results of the present study are similar to BSAF values calculated from literature values for infaunal deposit feeders, scavengers, filter feeders and benthically-coupled fish (Tracey and Hansen, 1996), where BSAFs for PAHs were uniformly lower (mean 0.34) than PCB (1.03) or pesticide (1.36) classes. In another study of Narragansett Bay species (SAIC, 1995b), these values were 0.27, 1.57 and 1.62, respectively.

Although only one species was collected in this study, the comparability of BSAF results for literature to multiple receptor groups demonstrates confidence in the calculated values. That is, the similarity in BSAFs for PCBs and pesticides across species and studies demonstrate that varying habitat of a target receptor, including epibenthic/filter feeders (ribbed mussels), infaunal (clams), epibenthic scavengers (lobsters), and epibenthic predators (fish) does not alter the bioavailability of organic chemicals in the sediments. PAHs, in contrast, appear to have lower bioavailability

than is typically observed, which is perhaps due to the form of PAHs in sediment at the Raymark site (perhaps highly weathered or combusted). Still, the functional contaminant exposure pathways among target receptors are similar, allowing the use of a single exposure pathway model to predict the ultimate fate (*i.e.*, tissue accumulation) of organic contaminants. The use of predicted residue concentrations will be discussed in Section 6.3.

6.2.1.2. Analysis of Metals Bioaccumulation

The bioavailability of metals for target receptors was assessed through Tissue Residue - Exposure Relationships analysis for calculation of Bioaccumulation Factors (BAFs). Whereas BSAF factors are based on Equilibrium Partitioning (EqP) theory, *i.e.*, non-ionic organic contaminants are assumed to be at steady-state between the carbon-normalized sediment concentration and the lipid-normalized tissue concentration (DiToro, *et al.*, 1991), the BAF model simply assumes that the sample ratio of tissue to sediment concentration is a measure of bioavailability.

Tissue Residue - Exposure Relationships. As discussed for organics, the degree of similarity of exposure to sediment associated contaminants was assessed for target receptors. A comparison of four inorganic contaminants (copper, mercury, lead, and zinc) in ribbed mussels with sediment concentrations are shown in Figure 6.2-3. Regression analyses were performed for four metal CoC-receptor pairings in order to elucidate possible correlations between tissue residues and metals exposure. Positive correlations were apparent between metals and tissue concentrations. First, a positive relationship was apparent between increasing copper concentrations in sediment with respective tissue concentrations ($r^2 = 0.59$). A similar relationship was true for lead ($r^2 = 0.50$). A very weak correlation was observed between mercury concentrations in sediments and respective receptor residue concentrations ($r^2 = 0.03$). No apparent relationship existed for sediment and tissue concentrations of zinc.

As the CoC-receptor correlation analysis for metals did not always produce meaningful relationships, the approach of Bioaccumulation Factors (BAFs) modeling was taken, which involves calculating the average ratio of CoC tissue residues from bivalves relative to sediment concentrations at Raymark stations. Because factors governing differential metals bioaccumulation among species are poorly understood relative to that for organic contaminants, the analyses were conducted on a metal-by-metal basis.

BAFs for metals were calculated for all CoC-ribbed mussel pairings. The mean of nine data points from Areas C and D was taken as the overall BAF for the analyte (Appendix D-3-4). Analyte-specific BAF values for metals fell into three groups relative to the propensity for accumulation into tissues:

High:	1)	cadmium (3.3)
Intermediate:	2)	mercury (0.93), copper (0.33), zinc (0.69)
Low:	3)	nickel (0.05), chromium (0.04), lead (0.04)

This pattern can be partly explained by the fact that the bioavailability of sediment-associated heavy metals is related to the concentration of iron oxides in sediment as well as insoluble sulfides (Bryan and Langston, 1992); sand-associated metals are expected to be more bioavailable than silt-associated metals for this reason. However, other differences in bioavailability may be explained by the chemical properties of the metals themselves. For example, the most bioavailable metal was cadmium. Because of relatively high solubility compared to most heavy metals, this metal exhibits extremely mobile behavior in aquatic systems, and in surface waters can undergo complex patterns of transformation including oxidation-reduction reactions, biotransformation, precipitation, and adsorption.

In the second group, mercury is well known to bioaccumulate in marine organisms, and has received considerable attention because of its toxicity relative to other metals (Wren et al., 1995). Remobilization of copper, mercury and zinc via resuspension and ingestion is the most probable exposure route of these CoCs to target receptors. Zinc is very persistent, and are generally removed from the environment by particle adsorption and subsequent settlement to sediments. The bioavailability of metals with intermediate BAFs is most likely affected by a variety of processes, ranging from dissolved-particulate partitioning to internal metabolic regulation; therefore, bioaccumulation is the result of the exposure to the metals in both water and sediment

Metals in the third group (chromium, nickel and lead) were the least bioavailable forms. Lead tends to be highly particle associated, while nickel and chromium are only moderately water-soluble. The particle-associated nature suggests that these metals are unlikely to be transported far from the source.

In summary, the observed bioavailability of metals in this study is generally consistent with the known behavior of metals with respect to mobility and solubility, and suggests that metal species will play an important role in controlling bioaccumulation in target receptors. Mean BAFs for CoC metals did not exceed unity, hence a lack of apparent biomagnification.

6.2.2. Tissue Residue-based Exposure Assessments

Site vs. Reference Tissue Concentration Ratios (TCRs) were employed to evaluate CoC tissue residues in ribbed mussels as an indicator or exposure. This analysis involves the comparison of receptor- and analyte-specific tissue body burdens from the Raymark study area stations against corresponding modeled data for the reference station (based on sediment chemistry at GM08; see SAIC, 1998). Due to data limitations, comparisons of site tissue concentrations against predicted reference

concentrations were made based on use of bioaccumulation factors for organics and metals (discussed in Sections 6.2.1.1 and 6.2.1.2, respectively). For organics data, tissue concentrations are typically normalized to the lipid content of the organism. In this study, tissue data were limited and the variance around measured lipid concentrations was small. Because variance in lipid concentrations was small between site and reference stations, it was determined that lipid normalization for organics data was not necessary.

Tissue Concentration Ratio (TCR) rankings for organic contaminants and metals in ribbed mussels from the Raymark study area are presented in Table 6.2-1; complete results for ribbed mussel TCRs are presented in Appendix D-4-1 (Note: reference station tissue concentrations are predicted from sediment concentrations (see Section 6.2) and are presented in Table Appendix D-4-1). Data were ranked according to the evaluation criteria described in Table 6.0-1. These were based on Total PCBs, 13 PAHs, a matrix-specific dioxin concentration (*i.e.*, 2,3,7,8-TCDD equivalent concentrations using fish TEFs, Appendix D-3-2), and seven anthropogenic metals.

Total PCB enrichment for ribbed mussels was 15- to 20-fold above reference concentrations at all stations in Area C. Five of the Area D stations were also ten to 40-fold above reference concentrations, including Stations D-1 through D-4 and Station D-6. Station D-5 had the highest Total PCB enrichment in Area D, measuring greater than 3200-fold above the reference. Station E-1 in Area E also had a TCR>40 for Total PCBs; predicted concentrations were 4000-fold higher than reference. The remaining three stations in Area E had TCRs for Total PCBs greater than ten, but less than 40. All three stations in Area F had Total PCB enrichment which was 48- to 58-fold above reference.

TCRs varied for PAH enrichment throughout the study areas (Table 6.2-1). Most PAHs for Area C had a baseline exposure. One exception was a low exposure (TCR>3) for phenanthrene at the three Area C Stations. TCR rankings for PAHs in Area D followed a similar pattern to those in Area C. At Stations D-1 and D-2, low CoC exposures were apparent for phenanthrene and pyrene. Baseline exposures were apparent for other PAHs at these two stations. Stations D-3 and D-4 had baseline conditions for all PAHs except phenanthrene, where a TCR > 3 was observed. TCR rankings for PAHs at Station D-5 were the highest of all Area D stations, where a greater than ten-fold enrichment occurred for six PAHs at Station D-5. Station D-6 had a low exposure for one PAH (Table 6.2-1).

In Area E, Station E-1 had, at most, low exposures for three PAHs (Table 6.2-1). Baseline TCRs were apparent for all PAHs at Stations E-2 and E-4, and five PAHs at Station E-3. In contrast, a greater than ten-fold enrichment occurred for seven PAHs at Station E-3, while low baseline exposure occurred for the remainder of PAHs at this station. TCRs at Area F indicated mostly intermediate (TCR>10) or low (TCR>3) exposure for the same seven PAHs at all three stations.

TCR rankings calculated for dioxins ranged from baseline to high exposures. Low exposure from dioxins was apparent for all stations in Areas C and F. TCR rankings for dioxins were generally low to negligible for stations in Area D, despite the fact that the greatest enrichment (TCR > 40) for dioxins in the entire study was observed at Station D-5. In Area E, an intermediate TCR ranking was apparent at Station E-1. Baseline exposures to dioxins were apparent in TCR rankings for the remainder of Area E Stations.

In summary, the highest and most extensive TCR rankings were for Total PCBs, particularly at Stations D-5 and E-1. TCR rankings for PAHs were generally highest throughout Area F (Selby Pond). Intermediate exposures for many PAHs were also apparent at Station E-3 adjacent to the baseball field, and at Station D-5, located in a wetland adjacent to the Housatonic River. Dioxin enrichment was also greater than 40-fold at Station D-5. TCR rankings were not calculated for pesticides since reference concentrations were not available. In contrast, baseline or low CoC exposure was apparent for metals throughout Areas C, D and E. Only Station F-2 on Selby Pond had slightly elevated exposures for metals (TCR > 3 for lead and zinc).

Overall exposure rankings for TCRs were mostly low and intermediate. Low exposures were observed in Areas C and D, with the exception of a high exposure at Station D-5. In Area E, two stations had an overall low exposure, while two had an overall intermediate risk. Intermediate risks were observed at all Area F stations. These results are carried forward into the overall risk ranking summary Table 6.5-1.

6.2.3. Tissue Residue-based Effects Assessment

In this section, effects of CoC residues in target species are evaluated by comparison of tissue concentrations against water quality-based benchmarks (Tissue Screening Concentration (TSC) HQs; Section 6.2.3.1), and thresholds for narcotic effects (Critical Body Residue (CBR) HQs, Section 6.2.3.2). In contrast to the TCR results which indicate degree of chemical exposure, the TSC and CBR evaluations address the likelihood of adverse impact on target receptors resulting from CoCs in their tissues.

6.2.3.1. Tissue Screening Concentration Assessments

The fundamental basis for Tissue Screening Concentration (TSC) Assessments is the assumption that the product of the Water Quality Criteria value for a given CoC and the Bioconcentration Factor (BCF, tissue-water concentration ratio) should provide an estimate of the tissue concentration which is also protective to the species. A recent study by Shepard (1998) involving a literature survey of paired residue/effects data for 100+ chemicals demonstrated that only 19% of the TSCs derived in the above manner were higher than tissue residues found to be associated with toxic effects (*i.e.*, the derived TSC was protective of aquatic life 81% of the time). CoCs for which the TSC approach was not protective primarily included PAHs; it was hypothesized that these

compounds exhibited enhanced toxicity through the activity of metabolic breakdown products which are not measured (Shepard, 1998). Exclusion of these compounds from the analysis improved the protectiveness of TSCs to a level comparable to that of water quality criteria (93%).

Given this first demonstration that TSCs can provide a level of protection equivalent to WQC, TSCs were adopted for this ERA to assess potential effects of tissue residue concentrations in target receptors. A Hazard Quotient approach was taken in this analysis, in which the measured tissue concentration is divided by the TSC effect concentration to calculate the TSC Hazard Quotient (TSC-HQ).

The analyte-specific residue concentrations presumed to be adverse to the organism (TSC effect concentrations) were calculated using the EPA WQSV for metals and taken as reported by Shepard (1998) for organics. These values are presented in Table 6.2-2. For the present study, CoC residues in target species, reported as a dry weight basis in Appendix Table A-3, were converted to wet weight (Table D-6-1 using sample-specific values that are reported in Appendix Table A-3) so as to allow direct comparison with the TSC benchmark. (Note: CoC tissue residues were predicted as described in Section 6.2 for stations from Areas E and F, Station D-5, and the reference station).

Rankings of TSC-HQ are presented in Table 6.2-3. Data were ranked according to the evaluation criteria described in Table 6.0-2. TSCs were evaluated based on Total PCBs, 14 PAHs, four pesticides, a matrix specific dioxin concentration (i.e., using fish Toxicity Equivalency Factors (TEFs), Appendix D-3-2), and seven metals.

TSC-HQs for Total PCBs in Area C were all below one, indicating no effects at these three locations. Area D stations also had TSC-HQs below one for all stations except Station D-5. This station had a TSC-HQ of 6.29 for Total PCBs, indicating a low effect on target receptors. Similarly, a low effect was also observed for one station in Area E; Station E-1 had a TSC-HQ of 7.81. The remainder of the stations in Area E indicated no effect for Total PCBs. The same was true for Area F stations and the reference.

PAHs were below effect levels for target receptors in all of the areas sampled and the reference site (Table 6.2-3). TSC-HQs were also below one for pesticides throughout the Raymark study site, as well. Matrix-specific dioxin concentrations produced TSC-HQs below one for all areas of the Raymark study site, including the reference.

It is apparent from the metals data that potential impacts on receptors, represented by ribbed mussels from the site as well as the reference location are caused primarily by this chemical class, particularly copper and zinc (Table 6.2-3). In Area C, TSC-HQs for copper and zinc ranged from 1.90 to 8.29 (Appendix D-6-2), indicating potentially low effects from these metals at all three sites. The same was

true for all Area D sites, low exposures were apparent for copper and zinc. In Area E, low effects were apparent for lead and zinc at all sites. TSC-HQs were below ten for copper at Stations E-2, E-3 and E-4, producing low effects. Intermediate effects were observed at Station E-1, with a TSC-HQ of 17.91 for copper. Area F stations had TSC-HQs ranging from below one to 49.04 (Appendix D-6-2). Station F-1 had low effects from copper, lead and zinc. TSC-HQs greater than 40 were observed for copper and TSC-HQs greater than 10 were observed for zinc at Station F-2. TSC-HQs above one were also present at this site for chromium, lead and nickel. Intermediate and low effects were apparent at Station F-3 for copper, lead and zinc. The reference station had elevated TSC-HQs for all metals with the exception of mercury; copper was particularly high (TSC-HQ = 101) and zinc was also elevated (TSC-HQ = 15).

Overall, these data indicate the highest potential impacts to target species are driven by exposure to copper in particular (TSC-HQ>40) in Area F (Selby Pond). Low effects were apparent in target species throughout all four areas of the Raymark study site. These low effects were also caused by metals, including copper, lead and zinc. In contrast, TSC-HQs for PAHs, PCBs and pesticides generally suggested a lack of residue-based adverse effects to the target species of the Raymark study site.

In summary, the magnitude of the TSC-HQs and spatial location suggest that metals are the primary source of tissue residue effects on target receptors across the study area as well as the reference locations. The overall TSC risk rankings showed mostly baseline conditions throughout Areas C, D, and E. One exception to this was an overall low ranking at Station E-1. Area F had overall risk rankings of baseline, low and intermediate for Stations F-1, F-3, and F-2, respectively. The overall TSC risk ranking for the reference stations was intermediate. These results are carried forward to the Tissue Residue Effects summary in Section 6.2.3.3 (Table 6.2-6).

6.2.3.2. Critical Body Residue Assessments

The previous analysis has addressed the potential impacts of individual CoCs when found at elevated concentrations within the tissues of target species. However, even when CoCs in target tissues are below the effects threshold for specific toxic action (*e.g.,* central nervous system, respiratory, digestive disruptions), the combined presence of the chemical mixture may still cause non-specific toxicity through a phenomenon called narcosis (McCarty *et al.*, 1992). Narcosis is a physiologically debilitating condition caused by CoC-related swelling of cell membranes. Critical Body Residues (CBRs) represent the concentrations at which narcotic effects on the species may occur. The utility of CBR analysis for this ERA is enhanced by the fact that the organic CoCs at the Raymark site (*i.e.,* PAHs, PCBs, pesticides, dioxins) generally fall into a class of compounds which can elicit a narcotic mode of toxic action (McCarty *et al.*, 1992). The CBR approach is also believed to apply to metals (McCarty and Mackay, 1993). Hence, narcosis theory does appear to be relevant to the various CoC classes considered in this analysis.

CBR values are expressed as the molar tissue concentrations (μMol/g dry wt) of CoCs measured in the tissue (*e.g.*, μg CoC/g dry weight divided by CoC molecular weight (μg CoC/μMol CoC)) and are compared against representative CBR benchmark values found in the literature (Table 6.2-4). Typically, CBR benchmarks are the Lowest Observed Adverse Effect (LOAEC) concentration after chronic (*e.g.*, > 3 weeks) exposure to a given CoC. Where CBR benchmarks are derived from acute exposures, however, the equivalent chronic critical body residue concentrations can be estimated from the acutely toxic CBR by applying the acute:chronic ratio of the respective water or sediment concentration toxicity relationship. For the present investigation, the total PAH LOAEC acute CBR value reported by Arnold and Biddinger (1995) was converted to a LOAEC chronic value assuming a 1:10 acute:chronic ratio. CBR benchmarks reported as wet tissue concentration are converted to a dry weight value assuming an average of 80% water content in the target species. Although it has been shown that tissue lipid concentration may increase the CBR benchmark by approximately 14% per each percent of lipid between 3% and 8% lipid concentration (McCarty and McKay, 1993), variation in lipid among target species was not considered in the present analyses.

To assess possible effects of tissue body burdens in the present ERA, measured tissue concentrations (calculated as μMol/g dry wt) of specific CoCs were divided by the available CBR benchmark found in the literature (Table 6.2-4) to derive CBR Hazard Quotients (CBR-HQs). In all, ten CBR benchmarks were obtained, including seven metals, Total PAHs, Total PCBs, and p,p'-DDE.

CBR-HQ rankings by chemical class are presented in Table 6.2-5. Data were ranked according to the evaluation criteria described in Table 6.0-2. CBR-HQ results reveal no effects to target receptors from CoCs in Areas C or D. A low potential for impacts due to copper was observed at Station E-1 near the baseball field, Station F-2 on Selby Pond and at the reference station (CBR-HQ < 3).

Overall risk rankings for CBR-HQs were baseline for all areas of the Raymark study site (Table 6.2-5). The CBR results are carried forward to the Tissue Residue Effects summary in Section 6.2.3.3 (Table 6.2-6).

6.2.3.3. Tissue Residue Effects Summary

Tissue Screening Concentration and Critical Body Residue Hazard Quotients have been evaluated in Section 6.2.2 to determine the probability of tissue residue-based effects for each station. When considered jointly, the sensitivity/accuracy of these metrics follows the pattern for ER-L and ER-M-based effects interpretation, in that the lower concentration metric (in this case, TSCs and ER-Ls) are expected to identify many more CoCs of possible toxicological significance (hence, greater sensitivity), but with less certainty of implied adverse effects than higher concentration metrics (*e.g.*, CBRs and ER-Ms). Conversely, the CBR metric represents a high accuracy/low sensitivity endpoint in that adverse field effects have been documented in some species

where CoC residues exceeded CBR benchmarks. However, the CBR value is typically derived from less sensitive endpoints (such as lethality), and acute:chronic ratios may be applied to acute measurements to predict possible chronic threshold values. This extrapolation results in increased uncertainty.

Tissue residue effects rankings are summarized in Table 6.2-6. The reported results are the maximum of the indicator-specific rankings; these results are carried forward to the weight of evidence summary presented in Section 6.6 (Table 6.5-1). As the data reveal, target receptors would be susceptible to possible adverse effects at stations in Area E, Area F and the reference, with the ranking determined by the TSC-HQ result. Specifically, Station E-1 near the baseball field, F-3 on Selby Pond, and the reference station are characterized by low effect levels. Station F-2 on Selby Pond had an intermediate tissue residue effect level.

The TCR values (discussed in Section 6.2.1) are an exposure indicator whereas TSC and CBR HQs are effects indicators. It is notable from the comparison of the two indicator types that the exposure indicator identified primarily organics whereas the effects indicators identified only metals as the primary CoCs of concern. The differences between the two findings can be attributed to the fact that the TCR analysis is reference-based, while the TSC and CBR analyses are compared to benchmarks. Although the organics are elevated in receptors of the Raymark study area, the residues are not likely to cause adverse effects due to low toxicity. In contrast, metals appear to be of concern at the site, but this concern also extends to the reference location due to an apparent ubiquitous distribution of the observed metals.

6.3. Trophic Transfer Effects

In sections below, trophic transfer of metals and organics to avian and mammalian receptors feeding on aquatic receptors are discussed. These relationships are presented in a framework intended to elucidate the essential operative transport and fate mechanisms that control chemical bioavailability and trophic transfer in the exposure pathway models for target receptors (outlined in Sections 3.4 and 3.5). Using these exposure models, the relative degree of CoC bioavailability in target receptors at Raymark stations versus the reference is discussed with respect to differences between species and habitat.

Adverse effects to avian and mammalian aquatic predators from the ingestion of contaminated food within the study area were assessed by comparison of prey concentrations and prey-derived CoC dosage to Toxicity Reference Values (TRV-Dose). A target species dosage model was employed to calculate uptake of CoCs as dependent upon exposure factors specific to the RoC (including size-dependent food consumption rate, foraging behavior, migratory behavior, and food preference) and compared to the NOAEL benchmark, discussed in Section 6.3.1. In Section 6.3.2, dietary No-Observable-Adverse-Effects-Level (NOAEL) benchmarks

were derived from data obtained primarily from laboratory tests and converted into values applicable to each receptor of concern (RoC) for the Raymark ERA, assuming that the laboratory-based benchmarks are applicable to similar species of similar body size at other sites (Sample et al., 1996).

6.3.1. Dose Calculations for Avian and Mammal Aquatic Receptors

The food-web exposure model was used to estimate the exposure of the receptor species through diet, expressed as a total daily dose. The receptors of concern in this study were black-crowned night herons (*Nycticorax nycticorax*) and raccoons (*Procyon lotor*). In the literature, most TRVs for terrestrial species are reported as the threshold daily dose to an individual. Estimating a site-specific dose (IR_T) allows for direct comparison of exposure estimates with TRVs. Contaminant body-burden data from the sampling of ribbed mussels plus water concentrations of CoCs, were used for input into the models. Incidental sediment ingestion was also used as an input variable where appropriate. The basic structure of the exposure model is:

$$\text{Equation 1:} \quad IR_{Total} = \sum_x IR_x = \sum_X [\sum_M [\frac{(C_{XM} \bullet IR_M) \bullet BF_{XM} \bullet HR}{BW}]]$$

Where:
IR_{TOTAL} = total ingestion rate of all contaminants (mg/kg bw/day dry weight)
IR_X = ingestion rate of contaminant X from all media
C_{XM} = concentration of CoC_X in medium$_M$ (mg/kg dry weight)
IR_M = ingestion rate of medium$_M$ (kg/day dry weight)
BF_{XM} = dietary bioavailability factor of CoC_X in medium$_M$ (percent)
HR = proportion of contaminated site relative to receptor species' home range (*i.e.*, exposure fraction) (unitless)
BW = body weight of receptor species (kg)

Ingestion Rate. Precise information on nutrition requirements and energetics of selected receptor species (Black-crowned night heron and raccoon) were not available from the literature. Instead, daily food and water intake rates have been estimated using an allometric equation based on their body weight in grams (Nagy, 1987). These equations for food ingestion, F, in units of grams dry weight per day, are as follows:

Black-crowned night heron	$FCR = 0.648 \times bw^{0.651}$	Equation 2
Raccoon	$FCR = 0.235 \times bw^{0.822}$	Equation 3

In addition, water ingestion, W, in units of liters per day were calculated from the generic models presented below:

Bird Water Ingestion $WIR = 0.059 \times bw^{0.67}$ Equation 4
Raccoon Water Ingestion $WIR = 0.099 \times bw^{0.9}$ Equation 5

Data on CoC concentrations in sediment, surface water, and key prey of the receptor species were incorporated into the model to estimate total chemical doses ingested according to their respective intake rates. The daily ingestion intake rates used in the dietary model are presented in Table 6.3-1a, which also details other exposure parameters used in equations above.

To account for ingestion of different food types by a given receptor, the ingestion dose of all prey items, plus sediment and water are summed. In this instance, only one prey item was used, hence the ingestion dose of prey will be calculated using the following equation (Equation 6):

$$\sum \left(C_{XM} \bullet I_{RM} \right) = \left(C_{fish} \bullet I_{fish} \right) + \left(C_{water} \bullet I_{water} \right) + \left(C_{se\,dim\,ent} \bullet I_{se\,dim\,ent} \right)$$

Black-crowned night herons are opportunistic feeders that consume a variety of aquatic species, and even small terrestrial mammals. Table 6.3-1b presents information on the composition of their diet, as reported by NOAA (1998a). It is apparent from Table 6.3-1b that this avian receptor has a number of food sources, however only one prey item was accounted for in this study due to a lack of data. Hence, ingestion rates of the single measured prey species were elevated to account for the unsampled items in the heron diet.

To estimate dry weight dietary exposure to the black-crowned night heron, ribbed mussels were collected from appropriate habitats. Ribbed mussels were used as a surrogate food source for fish in the diets of heron and raccoon receptors. The diet and feeding behavior of the herons suggests that incidental sediment ingestion does occur and therefore may be a significant exposure pathway (Beyer, pers. comm., 1995; Ohlendorf, pers. comm., 1995). Sediment ingestion was assumed to be equivalent to 5% of the total dry weight dietary intake. Also, the herons were estimated to consume 0.05 L of water per day based on their body size (Equation 4). Total concentrations of CoCs in surface water were used to estimate the dose for this component for the food-web model.

The diet and feeding behavior of raccoons is remarkably similar to that of herons, in that fish, crustaceans and insects are primary foods (U.S. EPA, 1993e) and incidental sediment ingestion does occur (Beyer, 1994). Dietary fractions for this species are reported in Table 6.3-1a and a summary of food consumption parameters are found in Table 6.3-1b. Again, it is noted that this receptor has a number of food sources, but only one prey species was used because data were lacking for others.

Bioavailability Factors. To account for differences in bioavailability of CoCs, a dietary bioavailability factor (BF) was applied for particular CoCs to adjust the estimated total daily dose. Dietary studies in which the dose was administered in the food source were targeted. Avian studies cited by Ammerman *et al.* (1995) found that 44% of copper and only 61% of zinc in plant food sources was absorbed by chickens. Using primarily animal protein sources, bioavailability of copper and zinc in chickens increased to 65% and 85%, respectively. For this assessment, the latter copper value was assumed for herons. For all other CoCs, the maximum assimilation in birds (85%) was assumed for the bioavailability factor (Bf_{XM}). The same was true for raccoons, in which 85% assimilation was assumed for all CoCs.

Home Range. The nearest black-crowned night heron colony is about 3.5 miles (5.6 km) from the Raymark facility. This species has been observed foraging in the tidal areas within 1.9 miles (3 km) of the facility, and along Middle and Upper Ferry Creek. Since information pertaining to home range and feeding territory were not available form the literature, assumptions were made regarding habitat use for the food-web model. Although it is generally accepted that black-crowned night herons defend a feeding territory, no information was available on territory size, making it difficult to arrive at a home-range exposure factor (HR) for the food-web model. With regard to wading birds, the size of the feeding territory depends on the bird's ability to defend it, which is positively correlated with body size. Territory size is also dependent on prey distribution, dictating the size of the are a bird must defend to obtain adequate food in an energy-efficient manner (Kushlan 1978). Consequently, the feeding territory of herons depends upon the physical conditions of the habitat. Black-crowned night herons will return to the same area to feed (Parsons, pers. comm., 1995). Due to their body size and site fidelity, it was assumed that the birds spent 100% of their time feeding in these areas. Accordingly, a home-range (HR) exposure factor of 1.0 was used in the food-web model.

A raccoon's home range is dependant upon its sex and age, habitat, food sources, and the season (Sanderson, 1987). It's most common home range appears to be a few hundred hectares, although values from a few hectares to more than a few thousand hectares have been reported. Winter ranges are smaller than ranges at other times of the year for both male and female raccoons, however, home ranges of males are larger than those of females, while the home range of females with young is restricted. Thus, it was realistic to assume that the raccoon spends up to 100% of its time foraging in the area of interest.

Body weight. For body weights of avian receptors, the maximum weights reported in U.S. EPA (1993) were used. For the raccoon, the average adult body weight was used. Both heron and raccoon data represent mean values for both males and females.

6.3.2. Toxicity Reference Values for Avian and Mammalian Aquatic Receptors

Toxicity Reference Values. The literature was reviewed for TRVs for birds and mammals for all CoCs at the Raymark facility. NOELs and LOELs were obtained from the primary literature, EPA review documents, and on-line database (IRIS). Tables 6.3-2a and 6.3-2b for herons and raccoons, respectively, present the TRVs used as benchmarks in the food-web model. These TRVs are expressed as daily dry weight doses of contaminants normalized to the body weight of the test species. Values were not available for all CoCs. NOELs were available for many, but not all, CoCs. For mercury, an avian LOEL was used with a one-half extrapolation factor (from U.S. EPA, 1993e) to arrive at a NOEL value. A comparison of all other LOEL-to-NOEL extrapolation values found that half the ratios are less than a factor of 3 (U.S. EPA, date unknown). A factor of one-tenth is used here to conservatively extrapolate from LOEL to NOEL levels (except mercury). Data are rarely available for the wildlife species of interest, and most often must be extrapolated from other species (*e.g.,* chicken, mallard). Because of this, TRVs for raccoon were scaled linearly to the body weight of the test species.

6.3.3. Assessment of Adverse Effects to Black-Crowned Night Herons

Calculated CoC doses to herons were compared to TRVs (TRV-HQ) to estimate adverse effects for this avian target species. Dosage estimates on prey (Appendix Table D-5-1) and sediment (Appendix Table D-5-2) ingestion were summed to obtain total assimilated exposure to CoCs (Appendix Table D-5-3). Station- and analyte-specific TRV-HQs for Black-crowned night herons consuming prey in the Raymark study area are presented in Appendix Table D-5-3. A visual summary of this information is presented in Table 6.3-3a, following the ranking strategy presented in Table 6.0-1.

Organics. TRV-HQs for Total PCBs summarized in Table 6.3-3a suggest baseline exposures for heron consuming prey for all stations in Area C, and most stations in Area D, Area F, as well as the reference station. Ingestion of prey at Station D-5 and at E-1 produced low exposure rankings for Total PCBs. TRV-HQs were 6.57 and 8.66, respectively (see Appendix Table D-5-3). However, baseline exposure to herons from Total PCBs occurred at the remainder of the Area D and Area E Stations.

Baseline exposure rankings for PAHs occurred across all four areas and the reference.

For pesticides, baseline TRV-HQs occurred throughout Areas C and E. In Area D, a low exposure for p,p'-DDD was apparent at Stations D-5 (TRV-HQ = 2.75). In Area F (Stations F-1 and F-3), low exposure rankings for p,p'-DDD were apparent; TRV-HQs were 1.71 and 1.76, respectively. In addition a low exposure ranking for p,p'-DDE was apparent at Station F-3 (TRV-HQ = 1.45).

Calculated TRV-HQs for dioxins were baseline throughout all of the Raymark study areas.

To summarize, overall potential risk to herons consuming prey from the Raymark study area from organic CoCs ranged between baseline and low rankings. PAHS and dioxin were not contributors to low risk rankings in any area as baseline exposures were apparent at all stations. Low potential risks attributable to PCBs at Stations D-5 and E-1, to p,p'-DDD at Stations D-5, F-1, and F-3, and p,p'-DDE at Station F-3 were apparent.

Metals. TRV-HQs for metals were generally baseline or low throughout the four study areas of Raymark. Chromium exposure rankings were baseline with the exception of Station F-2 and the reference station were low rankings were apparent. TRV-HQs were 1.9 and 1.12, respectively. Exposure to lead at F-2 produced a low risk ranking (TRV-HQ = 1.36) although rankings for lead at all other stations was baseline. Low mercury rankings were apparent at Stations E-1 (TRV-HQ = 1.38) and the reference station (TRV-HQ = 1.89) and low zinc rankings were assigned to Stations F-2 and F-3 (TRV-HQs = 3.33 and 1.28, respectively).

In summary, the overall ranking of modeled exposure to Black crowned night heron generally suggests baseline to low adverse effects across the site (Table 6.3-3a). Overall rankings were baseline, with low potential exposure at D-5, E-1, F-2, F-3, and at the reference station. The above effects analysis for CoC exposure to the heron was based on the NOAEL benchmark, which is a factor of ten below the Lowest Observable Adverse Effect Level (LOAEL) benchmark typically representing a chronic or sub-chronic (non-lethal) endpoint. In addition, the assumption that the heron would feed exclusively on prey species from Raymark is highly improbable. Hence, it is likely that the calculated HQs for CoCs in modeled exposure species of Raymark study area do not actually represent a high probability of adverse effect to the avian aquatic predators such as the Black-crowned night heron. The station- and CoC-specific rankings of potential adverse effects in Table 6.3-3a are brought forward into Table 6.3-4 to provide a comparison with effects on the mammal receptor, discussed in the following section.

6.3.4. Adverse Effects to Raccoons

Calculated CoC doses to raccoons were compared to TRVs (TRV-HQ) to estimate adverse effects for this mammalian target species. Dosage estimates were based on the summation of assimilated prey (Appendix Table D-5-1) and sediment CoCs (Appendix Table D-5-2) to obtain the total assimilated exposure to CoCs (Appendix Table D-5-4). Station- and analyte-specific TRV-HQs for raccoons consuming prey in the Raymark study area are presented in Appendix Table D-5-4. A summary of receptor specific Hazard Quotient rankings for raccoons are presented in Table 6.3-3b, following the ranking strategy presented in Table 6.0-1.

Organics. TRV-HQs calculated for exposure to Total PCBs suggest baseline exposures for raccoons consuming prey in Area C and at the reference station (Table 6.3-3b). For raccoons in Areas D and E, ingestion of prey at Station D-5 and E-1 produced low exposure risks (TRV-HQ = 7.01 and 9.71, respectively) from Total PCBs (Appendix Table D-5-4). Baseline exposures to Total PCBs were apparent at all of the remaining stations in Areas D and E.

Potential effects from PAH exposure were not apparent for raccoons at any locations in the study area. The same was true for pesticides and dioxins (Table 6.3-3b).

Metals. TRV-HQs for metals in Areas C and D showed baseline exposure for all metal CoCs except for copper and mercury. Raccoons had low potential effects due to copper exposure at Station F-2 in Area F and at the reference station (TRV-HQs were 1.16 and 1.01, respectively. Potential effects due to mercury exposure were apparent at Station E-1 (TRV-HQ = 1.19) in Area E and at the reference station (TRV-HQ = 1.62).

Generally, overall potential COC effects to raccoons were baseline. Low potential effects were apparent at E-1 and the reference station due largely to copper and mercury exposure. The station- and CoC-specific rankings of potential adverse effects in Table 6.3-3b are brought forward into Table 6.3-4 to provide a comparison of the effects on the avian receptor, discussed in the previous section. The above effects analysis for CoC exposure to the raccoon were based on the NOAEL benchmark, which is a factor of ten below the Lowest Observable Adverse Effect Level (LOAEL) benchmark typically representing a chronic or sub-chronic (non-lethal) endpoint. In addition, the assumption that the raccoon would feed exclusively on prey species from Raymark is highly improbable. Hence, it is likely that the calculated HQs for CoCs in prey species of the Raymark study area do not actually represent a high probability of adverse effects to mammalian aquatic predators such as the raccoon.

6.4. Analysis of Toxicity *versus* CoC Concentrations

This section evaluates the relationship between the sediment CoC concentrations and sediment toxicity. The analysis of results focuses on elucidation of potential exposure-response relationships for in-place sediment (*e.g.*, bedded sediment). For this assessment, the measurement endpoint evaluated was the toxicity of bulk sediments to amphipod survival.

Toxicity occurs when CoCs in the environment become bioavailable above concentrations which cannot be physiologically managed by the organism. In the present investigation, tissue residues of laboratory bioassay species were not measured, hence the primary evidence to ascertain apparent toxicity and identify the potentially responsible CoCs was to examine the degree to which the measured CoCs

in sediments exceeded a known benchmark or criteria, and assess the strength of relationships between the observed toxicity and the exposure concentrations measured in sediments.

As discussed in Section 4.3, divalent metal bioavailability may be predicted from sediment SEM and AVS relationships. In contrast, the bioavailability of organic CoCs is believed to be controlled by the partitioning between the organic carbon fraction of sediment/porewater and the lipid fraction of tissue. Hence potential CoC exposure - toxicity relationships are considered in relation to the bulk sediment concentration as well as the Total Organic Carbon (TOC) normalized concentration, with the inherent assumption that high TOC sediments will have lower CoC bioavailability than low TOC sediments with the same unit CoC concentration. In addition, sediment TOC measurements are used in lieu of dissolved organic carbon data (not measured in this study) when assessing bioavailability of CoCs in porewater preparations with the assumption that sediment TOC concentrations should be positively and linearly correlated.

A general assumption is that some correlation exists between the degree of mortality in a toxicity test and the extent of potential effects on the benthic community, however the exact relationship is not known. Lacking specific criteria, reasonable cutoff values are applied. For example, the 80% survival cutoff for the amphipod bulk sediment toxicity test is derived from a statistical evaluation of published toxicity results (Thursby *et al.*, 1997) to determine the minimum degree of toxicity that represents an 80% probability finding of a statistically significant reduction relative to controls. It is noted that this cutoff is merely a statistical threshold, as the ecological relevance of 20% reduction in survival is not known. Recently, Scott (1998) has demonstrated a correlation between survival of *Ampelisca* in sediment toxicity tests with benthic community condition. A 60% threshold was selected due to an accurate prediction of degraded benthic community and this degree of toxicity. Lacking additional guidance, other breakpoints are selected based on best professional judgement. Still, the risk manager is encouraged to consider the sum of the evaluations and conclusions made based on the available data.

The amphipod toxicity response to bedded sediment was evaluated by comparison of relationships between survival versus 1) bioavailable metals (related to SEM and AVS concentrations); 2) sediment ER-M HQs for organic contaminants and metals; and 3) Porewater WQC-SA HQs for representative organic contaminants (Total PAHs, Total PCBs, pesticides and dioxins).

Toxicity versus sediment concentrations for organic contaminants and metals. Calculated HQs for concentrations of organic contaminants and metals in sediments were plotted versus amphipod survival to examine the data for possible relationships. Patterns observed for amphipod survival versus sediment ER-M HQs for organics (Figure 6.4-1A), were generally not suggestive of exposure-response relationships. The highest TOC-normalized concentration of Total PAH was detected at Station D-3,

where reduced amphipod survival was observed (Figure 6.4-1B). Normalized concentrations of Total PCBs were elevated at Stations D-5 and E-1, and somewhat reduced amphipod survival was observed. The same pattern was true for TOC-normalized dioxins. Hence, it does not appear that reduced survival in amphipods is related to sediment-based measures of Total PAHs, Total PCBs, p,p'-DDD or dioxins. Thus, despite the fact that reduced amphipod survival was observed at six stations, concentrations of CoCs in sediments did not appear to be related to the observed effects.

ER-M HQs for four metals were plotted versus amphipod survival to examine the data for possible exposure-response relationships. Increasing sediment HQs for copper, lead, mercury and zinc did not appear to adversely affect amphipod survival (Figure 6.4-2).

Bioavailable metals related to SEM and AVS relationships. Relationships between amphipod toxicity and three measures of metal bioavailability are presented in Figure 6.4-3. Because of the volatility of AVS in the presence of oxygen, and hence the possibility that some AVS could be lost during sampling or analysis, the relationship between amphipod toxicity vs. SEM metal concentration was investigated. The resulting relationship (Figure 6.4-2A) indicated no apparent correlation between amphipod survival and SEM metals concentration, suggesting that metals at Station C-3 and D-6 were not responsible for observed toxicity if AVS was lost prior to or during laboratory exposures. Further inspection of the data using the difference of SEM and AVS (SEM-AVS) as the indicator of metal bioavailability again suggested no apparent exposure-response relationship, although survival at Station C-3 was below the probable impact level where the SEM-AVS result indicated potential metal bioavailability (Figure 6.4-1B). The difference between SEM and AVS was also normalized to the organic carbon fraction of the sediments (Figure 6.4-1C) so as to account for non-AVS compounds which might also bind metals into a non-bioavailable form. Increased toxicity was apparent where the normalized concentration exceeded zero, yet the toxicity to amphipods at Station D-6 was still not explained by this data normalization.

Toxicity versus porewater concentrations for organic contaminants and metals. Patterns observed for amphipod survival versus porewater HQs for organics (Figure 6.4-4) were also not suggestive of exposure-response relationships. The same was true for increasing copper, lead and zinc HQs in porewater (Figure 6.4-5); examination of these individual relationships between metals concentrations and toxicity still did not produce an explanation for decreased toxicity, particularly at Stations C-3 and D-6.

6.5. Risk Synthesis

The interpretation of ecological risk in this assessment is based on a weight of evidence approach. The weight of evidence is in turn based on the analysis of

exposure and effects data, as represented by the endpoints discussed in the exposure (Section 4.0) and effects (Section 5.0) sections of this ERA, as well as in risk characterization (Sections 6.1 through 6.4). The individual weights of evidence were interpreted and summarized using semi-quantitative ranking schemes so as to allow their inclusion into an analysis of the overall risk indicated for each of the primary weight of evidence categories. In Sections 6.5.1 and 6.5.2, below, the process of synthesizing information obtained on individual indicators and translating the result into an Exposure/Effects Weight of Evidence (WoE) ranking is presented. The primary exposure-based WoEs for Bedded Sediment are SEM Bioavailability and Porewater Sediment Hazard Quotients. Tissue Concentration Ratios of receptors allow evaluation of CoC bioconcentration in organism tissue. The primary effects-based WoE are Sediment Toxicity, Tissue Residue Effects, and Trophic Transfer impacts to Avian and Mammalian Predators consuming prey in the Raymark study area.

Results of the evaluations of the WoE data are presented in exposure and effect WoE summary tables in Sections 6.5.1 and 6.5.2, respectively. In Section 6.5.3, the findings of exposure and effects WoE are evaluated jointly in order to interpret the overall potential for adverse ecological risks by sampling station.

6.5.1. Exposure-Based Weight of Evidence

Exposure-based weights of evidence include Bedded Sediment Exposure (based on HQs for CoC sediment and porewater contaminants, SEM metal bioavailability) and CoC residues in target species relative to reference as assessed through Tissue Concentration Ratios (TCRs).

Bedded Sediment Exposure. Chemical concentrations of CoCs measured in sediment-associated media are compared against benchmarks to elucidate potential adverse effects on target species from exposure to contaminant concentrations in surface sediments. The weight of evidence for indicators of chemical exposure in bedded sediments suggest the highest probability of adverse exposures occur at Station D-3 in a wetland adjacent to the Housatonic River (Table 6.5-1). It is noted that the primary media/CoCs driving this risk are sediment PAHs (see Table 6.1-5 and supporting Table 6.1-1). The bedded sediment exposure ranking is also driven by elevated PAHs and metals concentrations, particularly copper in sediment porewaters (Table 6.1-4). The other divalent metals (cadmium, lead, nickel and zinc), in contrast, do not contribute substantially to the overall exposure ranking (Table 6.1-5).

Bioconcentration. Bioconcentration of CoCs in site receptors was assessed through Tissue Concentration Ratios calculated as the station-specific residue concentration at the site compared to the reference location for each CoC-receptor pair (Section 6.2). The metric is intended to elucidate those CoCs and receptors which are chemically enriched at the site relative to regional background conditions. Hence, it is principally an indicator of chemical exposure but does not predict effects.

CoC residues suggest mostly low or intermediate exposure of CoCs for ribbed mussel receptor species at the Raymark stations (Table 6.5-1). Ten stations had an overall low Tissue Concentration Ratio ranking, primarily driven by Total PCBs (Table 6.2-1). Five stations (E-1, E-3, F-1 to F-3) had intermediate exposure, primarily due to Total PCBs with some PAHs. One station had a high TCR exposure ranking, mostly due to Total PCBs and dioxins. CoC residues for Total PCBs were always more than ten-fold higher than reference values, but in many cases higher than forty-fold. Two sites (D-5 and E-1) had predicted CoC residues greater than 3000-fold above the reference site.

Exposure Ranking. Stations were ranked according to overall exposure and these rankings are presented in Table 6.5-1. Low exposures ("L") were apparent in Area C stations. Four stations in Area D (D-1, D-2, D-4, and D-6) also had overall low exposures to CoCs, as well as Station E-4. High chemical exposures ("H") were apparent for two stations in the Raymark study area, Stations D-3 and D-5. All other stations had intermediate ("I") exposures for CoCs (Table 6.5-1).

6.5.2. Effects-based Weight of Evidence Summary

<u>Sediment Toxicity</u>. In this ERA, the sediment bioassays with the amphipod, *Ampelisca* were used to assess possible impacts from in-place sediments. Laboratory toxicity results generally indicated some degree of sediment toxicity to amphipods throughout the Raymark study site.

The overall station-specific laboratory toxicity rankings are summarized in Table 6.5-1. High toxicity was observed at two stations (C-3 and D-6), while intermediate toxicity occurred at six stations (C-1, C-2, D-2, D-3, E-4 and F-1). Seven stations (D-1, D-4, D-5, E-1, E-2, E-3 and F-3) had low toxicity to amphipods (including the reference), and one remaining station was non-toxic to amphipods (F-2). As noted in Section 4, exposure response relationships explaining the observed toxicity were not readily evident.

<u>Tissue Residue Effects</u>. As discussed in Section 6.2, possible impacts of CoC residues on target species were assessed separately through Tissue Screening Concentration (TSC) and Critical Body Residue (CBR) Hazard Quotients. A summary of these tissue residue-based effects results is presented in Table 6.5-1. The tissue residue effects rankings were mostly baseline. Only four stations were above baseline, including two stations with low effects (E-1 and F-3) due to copper and two stations with intermediate effects (F-2 and reference) due to copper and zinc.

<u>Trophic Transfer Effects.</u> Trophic transfer effects parameters, summarized in Table 6.5-1 include avian and mammalian predator effects. The food web modeling for avian and mammalian aquatic predators assumed that Black-crowned night herons and raccoons were feeding maximally on the most contaminated of prey items available at a given station. Despite the conservative assumptions employed, five stations had no

higher than low effects rankings and the remaining stations were assigned to the baseline effects category. The low effects rankings were due to PCBs, pesticides, and chromium, lead, mercury, and zinc exposures to herons and to PCBs, copper, and zinc exposures to raccoons.

Ecological Effects Ranking. Overall effects to biological receptors from CoCs are summarized in Table 6.5-1. One station in the Raymark study area had a baseline ("B") effect ranking (Station E-3). Seven stations had low ("L") effects, and these were Stations D-1, D-4, D-5, E-1 - E-3, and F-3. Overall high ("H") effects were observed at Stations C-3 and D-6. The eight remaining stations had overall intermediate ("I") effects.

6.5.3. Synthesis of Exposure and Effects Weights of Evidence

Discussion of each of the weights of evidence and applicable exposure-response relationships has been presented in the previous sections. The focus of this section is to elucidate concordance among exposure-based and effects-based weights of evidence, in order to characterize overall potential risk for the Raymark study area.

The synthesis of risk is supported by the information presented in the Exposure and Effects summary table, as well as equally important evaluations of the strength of exposure-response relationships and/or presence of confounding factors which could artificially mask or enhance perceived risks. The Exposure/Effects rankings for stations are equal to the maximum of individual WoE rankings. Whereas the overall probability of adverse Exposure/Effect (E/E) WoE is based on the degree of agreement between exposure and response WoE summaries, as follows:

Baseline risk is defined as the probability of adverse exposure and/or ecological effects equivalent to that from contamination and other environmental conditions not associated with the site. From a ranking perspective, a baseline risk ranking required that only Baseline (B) ranking for E/E WoE summaries is observed.

A *Low* probability of ecological risks suggests possible, but minimal impacts based on some of the exposure or effects-based weights of evidence, while impacts are undetectable by the majority of exposure and effects-based weights of evidence. Conditions of low risk probability typically lack demonstrable exposure-response relationships. A Low Risk ranking requires that no greater than Low (L) ranking for E/E WoE summaries be observed or Intermediate (I) ranking for one WoE summary and no greater than Low (L) ranking for the other WoE summary.

An *Intermediate* probability of ecological risk occurs for site conditions falling between high and low probabilities of risk. As such, the intermediate risk probability condition is typically characterized by multiple exposure or effects weights of evidence suggesting that measurable exposure or effects, but not

both, are occurring at the site. Typically, quantitative exposure-response relationships are lacking. Intermediate risk probability may also be indicated if the spatial extent of apparent impact is highly localized (*e.g.*, a single station), or if the impact occurs for periods of very limited duration. To assign an intermediate Risk ranking, either Intermediate (I) risk ranking is observed for both E/E WoE summaries, or High (H) ranking for one E/E WoE summary and Low (L) ranking or Intermediate (I) for the other E/E WoE summary.

Conditions indicating *High* probability of ecological risk occur when numerous weights of evidence suggest pronounced contaminant exposure and effects, the spatial extent of apparent impact is great, the impact is likely to be persistent over long periods of time, and the available data support demonstrable exposure-response relationships. A High Risk ranking requires that a High (H) ranking be observed for both WoE summaries or a High ranking for one WoE and Intermediate ranking for the other WoE.

This type of ranking scheme is intended only as a qualitative tool to provide definition and uniformity for the description of potential risks as discussed in the following section. The ranking approach is based on best professional judgement, since the "true" ecological risk of, for example, benchmark exceedence or observed toxicity, is not presently known. It is not intended to place rigorous boundaries on actions that risk managers may take with respect to the results of the study. Hence, the risk manager is encouraged to keep in mind the nature of the risk ranking approach when evaluating the general outcome of the risk assessment.

The summary of exposure-based and effects-based weights of evidence and characterization of potential risk for the Raymark Phase III Ecological Risk Assessment is presented in Table 6.5-1 and discussed by risk category, below.

High Risk Probability Stations. In the present investigation, only Station D-3 is categorized as a high risk station, given both high exposure and high effects rankings. In addition, some support for exposure-response relationships were observed given that toxicity was observed and PCB concentrations in sediment were well above ER-M thresholds.

Intermediate Risk Probability Stations. Stations for which the WoE demonstrate intermediate risks include Stations C-1, C-2, C-3, D-2, D-5, D-6, E-1 to E-4, F-1 to F-3, and the reference station. Multiple exposure- or effects-based weights of evidence were observed in the data, resulting in an intermediate Exposure and/or Effects ranking. However, quantitative exposure-response relationships were found to be lacking.

Low Risk Probability Stations. A low risk probability was indicated for the remaining Raymark stations (D-1 and D-4). Minimal impacts are suggested by

the majority of exposure and effects-based weights of evidence, and no exposure response relationships were evident.

Baseline Risk Probability Stations. Baseline risk was not assigned for any of the Raymark stations.

6.6. Risk Uncertainty

Uncertainty Factors Related to Weights of Evidence (WoE). The weight of evidence in this assessment is dependent upon analyses of exposure and effects data, and their integration into risk characterization determinations. The purpose of the uncertainty analysis is to identify the potential uncertainty sources as well as their possible relationship to the true degree of adverse exposure or effects as inferred from field measurements and laboratory tests used to support the individual WoE. Depending on the nature of the test endpoint or its method of interpretation, the uncertainty may tend to either over- or underestimate the true degree of adverse impacts (*e.g.*, "false positive" and "false negative" results, respectively).

For the present investigation, lists of potential uncertainties believed to be important for exposure and effects measurement endpoints are summarized in Tables 6.6-1 and 6.6-2, respectively, and are discussed in the following sections.

Uncertainties discussed in the exposure phase of this assessment (Section 4.3) included:

- Adequacy of CoC selection and contaminant behavioral characterization;

- Adequacy of fate and transport evaluations, including station selection, spatial (horizontal) and vertical (sediment layering) patterns, and sample representativeness;

- Adequacy of characterization of temporal/spatial variability in CoC distribution; and

- Reliability of exposure point estimation methods, including sampling methods for SEM and AVS and data utilization.

Uncertainties discussed in the effects assessment phase (Section 5.5) included:

- Adequacy of toxicity data, including comparability between test species and species present;

- Adequacy of toxicity testing methods;

- Appropriateness of chemical concentration benchmarks for tissue residues;

- Adequacy and availability of national criteria as benchmarks;

- Appropriateness of the selected bioassay species as surrogates for the indigenous community; and

- Lack of data for measurements of benthic community structure, bivalve condition indices, hematopoietic neoplasia effects, and PAH effects on fish (P450 enzyme activity).

These exposure and effects uncertainties compound one another, as exposure and effects data are integrated in the risk characterization. In addition to these uncertainties, there are additional uncertainties which have been identified during the risk characterization, including:

- Limited toxicological data for target receptor species;

- Incomplete knowledge of community ecology, including natural history (*e.g.,* size of feeding range and site use) of many species, receptor sensitivities to contaminants, and trophic transfer of CoCs, and natural changes and variability in biological/ecological systems; and

- Adequacy of bioaccumulation and toxicological models.

In the present ERA, tissue residues have been used as an indication of exposure and possible effects; however, their utility as weight of evidence in ecological risk assessments is currently limited since evidence linking ecological effects directly with contaminant concentrations in tissue is generally lacking. In addition, more complete understanding of bioaccumulation and trophic transfer is required to evaluate the role of tissue residues in the status of natural resources, and to provide data for evaluating risks to human health associated with seafood consumption.

It is also noted that the Tissue Concentration Ratio (TCR) results employed as a weight of evidence for exposure directly depend upon use of reference data, such that the quantitative evaluation of potential ecological risk at a given station is generally dependent upon the reference condition. Some elevated concentrations of metals were observed at the reference station (see discussion in Section 4.3), such that TCR ratios may be reduced and thus underestimate site-related risk.

The application of organic (BSAF) and inorganic (BAF) bioaccumulation models have several uncertainties. The BSAF model relies on an empirical assumption that porewater concentrations are in equilibrium with sediment concentrations. This may not be the case, especially at sites such as Raymark where CoC releases could vary tidally

over the short- and long-term. Uncertainty with BAF models (*e.g.*, species-specific bioaccumulation patterns for various metals) is highly site-specific and may vary among species.

Uncertainties associated with the calculated Hazard Quotients exist because they do not necessarily reflect all chemicals or activities of chemical mixtures. An additive approach to HQs was taken in order to integrate multiple contaminant effects, since information is very limited on the toxicity of simultaneous exposure to mixtures of contaminants. However, this estimation does not incorporate potential synergistic or antagonistic interactions among chemicals, nor does it encompass risks from chemicals which were not measured.

Given that Risk Characterization is a synthesis of findings from the Exposure and Effects Assessments, it follows that uncertainties associated with these components of the Risk Assessment can be nullifying, additive or even compounded. A prime example is in the application of Hazard Quotients, where the numerator and denominator each represent point concentrations with an unknown departure from the "true" concentration. Exposure-toxicity relationships suffer the same uncertainty; separate error in estimates of survival and exposure concentration, for example, can compound or obscure true dose-response relationships or falsely suggest others which are misleading or unfounded.

The weight of evidence approach to characterization of potential risk is effective in reducing uncertainty because the probability that multiple exposure and effects indicators could spuriously suggest risk (or lack of it) decreases as the number of indicators in agreement increases. However, this approach in fact only reduces uncertainty with respect to the location and magnitude of risk. It does not specifically address the ultimate source of this risk (*i.e.*, Raymark vs. other contaminant sources), nor does it address potential future use scenarios involving fundamentally different conditions or activities at the site. This uncertainty has not been addressed in the present study through the inclusion of multiple reference stations and the analysis of spatial trends in CoC distributions which might suggest alternative CoC sources.

Estimation of Uncertainty in Risk Designations. At the high probability risk station (D-3) contaminant exposure and effects relationships was not substantially demonstrated. The nature of the contamination, being in the sediment, suggests the impact that the risk is likely to be persistent over a long period of time, although the spatial extent of apparent impact may be limited as nearby stations do not display similar levels of risk. This suggests that the overall uncertainty of the risk designation is intermediate.

At the low risk stations (D-1 and D-4), the majority of exposure and effects-based weights of evidence were low, and few exposure-response relationships were observed between exposure and effects indicators. This suggests that the overall uncertainty of the risk designation is low.

Among the thirteen Raymark stations assigned an intermediate probability of ecological risk, the data suggests that measurable (and occasionally somewhat high) exposure and/or effects were occurring, but not generally as high as for the high risk station described above, but like the high risk station, quantitative exposure-response relationships are generally lacking. For all of the stations, elevated potential risks from CoC residues in target receptors was accompanied by similar degree of risk related to CoC exposure. The balance between exposure and effects indicators lends high confidence about conclusions of risk magnitude.

Implicit in the characterization of risks at the site is a qualitative comparison of exposures and effects observed in site samples in comparison to that observed for the reference location: while the majority of the lines of evidence are developed independently of reference condition (e.g., sediment and porewater HQs, sediment toxicity) others are entirely dependent (e.g., tissue concentration ratios). Reference areas in the vicinity of the Raymark site have been observed to have elevated CoC concentrations which are not believed to be site-related. Notable among the metals is copper, which at 660 μg/g in reference sediment, exceeds all but one of the site locations. Measured porewater concentrations of Cu, Ni, Ag and Zn were also higher at the reference location than any other site. The location of reference station GM08 being in the middle of the expansive Great Meadows Marsh, would appear to place it far way from point sources and the data is comparable to that found by NOAA (1998) during their evaluation of reference areas for the Phase II (Area A) Ecological Risk Assessment. Thus, the existence of regionally high CoC concentrations away from the site introduces considerable uncertainty as to the extent of incremental risks posed by site-related contaminants.